Machine Learning for Radiometer Calibration

Samuel Alan Kossoff Leeney

3rd Year PhD Candidate

With: Harry Bevins, Eloy de Lera Acedo, Will Handley, Rohan Patel, Kaan Artuc, Jiacong Zhu

REACH

UNIVERSITY OF CAMBRIDGE

erc European Research Council Established by the European Commission

UKRI

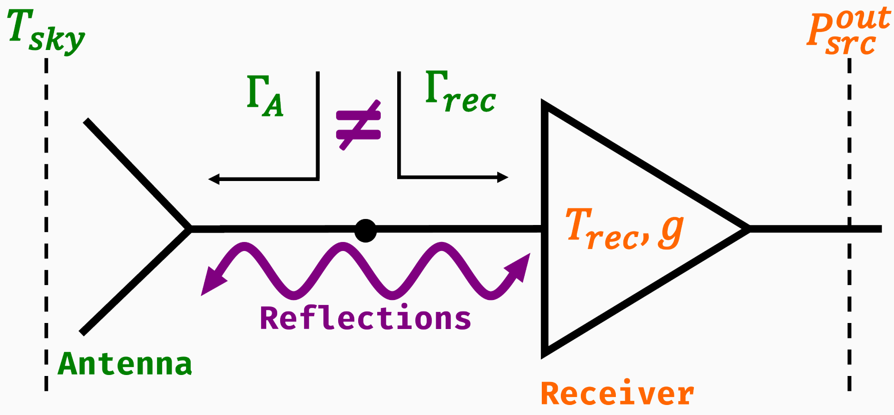


Figure 1: Schematic of the receiver system showing the signal path from antenna to digitizer

How to calibrate

Objective: Map input temperature to output power.

Key Factors:

- LNA introduces time-dependent gain, $g(t)$.
- Impedance mismatch adds noise (T_{rec}) to the system.

Link Output Power to Input Temperature:

$$P_{\mathrm{out}}^{\mathrm{src}} = gM \times (T_{\mathrm{in}}^{\mathrm{src}} + T_{\mathrm{rec}}) \qquad (1)$$

$$M = (1 - |\Gamma_{\mathrm{cal}}|^2) \left| \frac{\sqrt{1-|\Gamma_{\mathrm{rec}}|^2}}{1-\Gamma_{\mathrm{cal}}\Gamma_{\mathrm{rec}}} \right|^2$$

Note: All parameters above are frequency-dependent, but the notation has been simplified here and thereafter for convenience.

Dealing with reflections...

$$P_{\text{out}}^{\text{src}} = gM\left(T_{\text{in}}^{\text{src}} + T_{\text{rec}}\right) \quad (2)$$

Noise Parameter Equation:

$$P_{\text{out}}^{\text{src}} = gM\left(T_{\text{in}}^{\text{src}} + T_{\text{min}} + T_0 \frac{4R_N}{Z_0} \frac{|\Gamma_{\text{src}} - \Gamma_{\text{opt}}|^2}{(1-|\Gamma_{\text{src}}|^2)(1+|\Gamma_{\text{opt}}|^2)}\right) \quad (3)$$

Noise Wave Equation:

$$\begin{aligned} P_{\text{out}}^{\text{src}} = g\Bigg[& T_0 + T_{\text{unc}}|\Gamma_{\text{s}}|^2 \left|\frac{\sqrt{1-|\Gamma_{\text{rec}}|^2}}{1-\Gamma_{\text{s}}\Gamma_{\text{rec}}}\right|^2 \\ & + T_{\text{s}}(1-|\Gamma_{\text{s}}|^2)\left|\frac{\sqrt{1-|\Gamma_{\text{rec}}|^2}}{1-\Gamma_{\text{s}}\Gamma_{\text{rec}}}\right|^2 + T_{\text{cos}}\Re\left(\Gamma_{\text{s}}\frac{\sqrt{1-|\Gamma_{\text{rec}}|^2}}{1-\Gamma_{\text{s}}\Gamma_{\text{rec}}}\right) \\ & + T_{\text{sin}}\Im\left(\Gamma_{\text{s}}\frac{\sqrt{1-|\Gamma_{\text{rec}}|^2}}{1-\Gamma_{\text{s}}\Gamma_{\text{rec}}}\right)\Bigg] \end{aligned} \quad (4)$$

- g - Gain
- T_{min} - Minimum Noise Temperature
- R_N - Noise Resistance
- Γ_{opt} - Optimum Reflection Coefficient
- Γ_{s} - Source Reflection Coefficient
- Γ_{rec} - Receiver Reflection Coefficient
- T_{s} - Source Temperature
- $P_{\text{out}}^{\text{src}}$ - Power out
- $T_{\text{unc, cos, sin}}$ - Noise wave parameters
- T_0 reference temperature

Radiometer calibration with machine learning

ML calibration overview

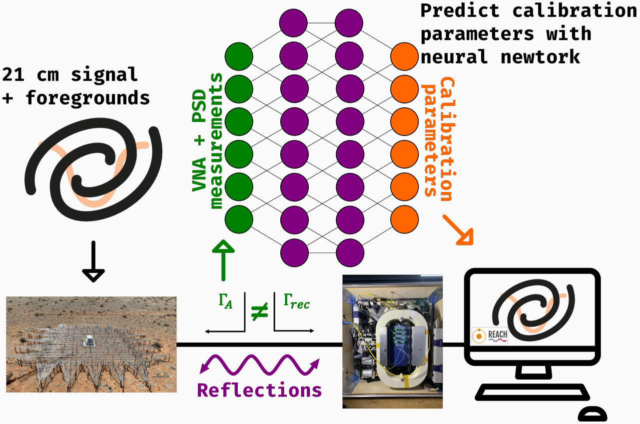


Figure 2: High-level overview of the machine learning-based calibration framework

Machine learning calibration steps

1. Define the Loss Function

Regress over measured power and predicted power.

$$\mathcal{L} = \frac{1}{n} \sum_{i=1}^{n} \left(\mathcal{P}_{\text{measured},i} - \mathcal{P}_{\text{pred},i} \right)^2 \tag{5}$$

2. Write Down the Equation for $\mathcal{P}_{\text{pred}}$

Using the noise wave formalism, relate $\mathcal{P}_{\text{pred}}$ to T_{src}.

$$\mathcal{P}_{\text{pred}} = g \cdot M \Bigg(T_{\text{in}}^{\text{src}} + T_{\text{min}} + T_0 \frac{4R_N}{Z_0} \frac{|\Gamma_{\text{src}} - \Gamma_{\text{opt}}|^2}{(1 - |\Gamma_{\text{src}}|^2)(1 + |\Gamma_{\text{opt}}|^2)} \Bigg) \tag{6}$$

3. Minimise the Loss Function

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} \mathcal{L}(\boldsymbol{\theta}) \tag{7}$$

parameter vector $\boldsymbol{\theta}$ includes all tunable parameters in the model:

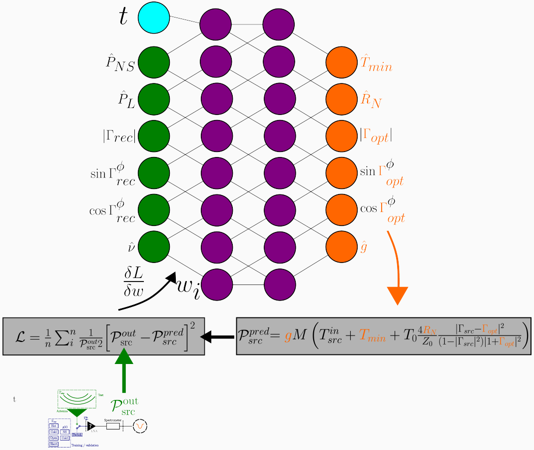
$$\boldsymbol{\theta} = \{g, T_{\text{min}}, R_N, \Gamma_{\text{opt}}^{\phi}, |\Gamma_{\text{opt}}|) \tag{8}$$

4. Rearrange and Predict (T_{src}) using $\boldsymbol{\theta}^*$

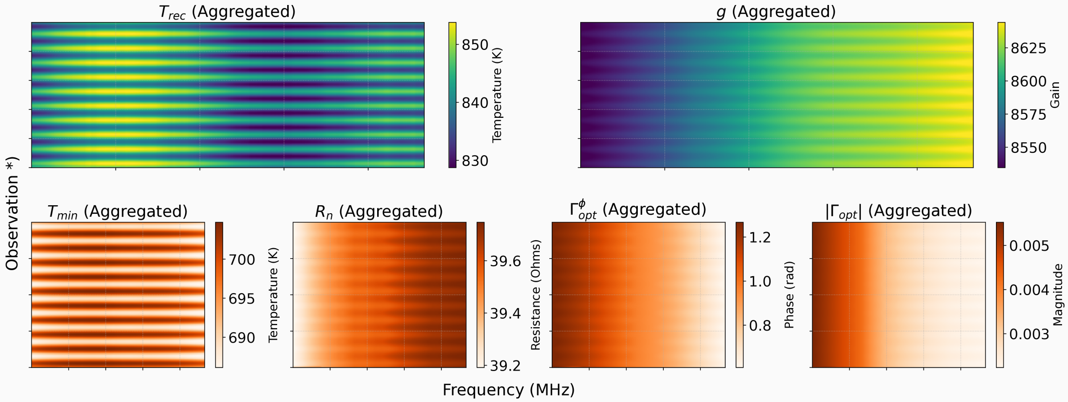
$$T_{\text{src}} = \frac{\mathcal{P}_{\text{pred}}}{g^* \cdot M} - \Bigg(T_{\text{min}}^* + T_0 \frac{4R_N^*}{Z_0} \frac{|\Gamma_{\text{src}} - \Gamma_{\text{opt}}^*|^2}{(1 - |\Gamma_{\text{src}}|^2)(1 + |\Gamma_{\text{opt}}^*|^2)} \Bigg) \tag{9}$$

Learning non-linear time-dependent system drift

Neural Network Time Evolution

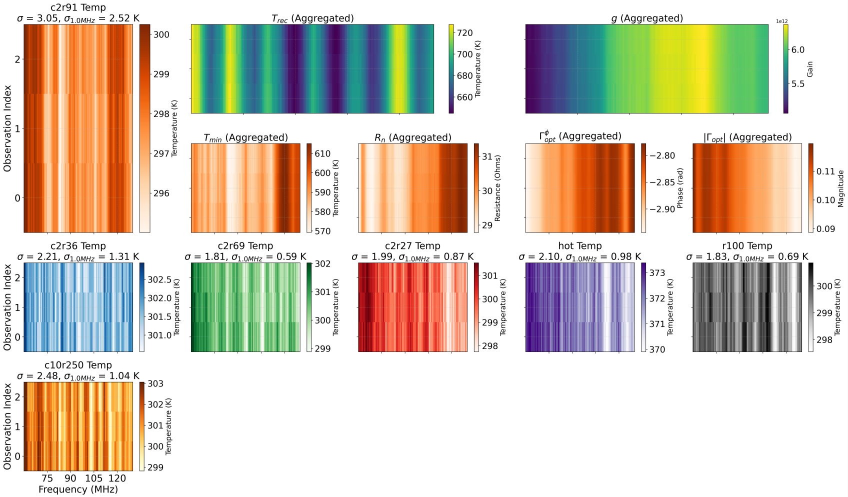


Inject system drift

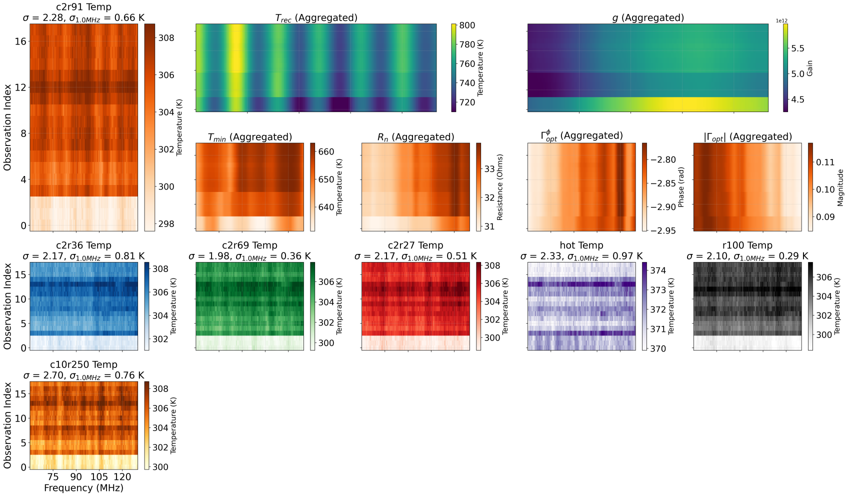


Inject a time varying sinusoid into T_{min} and predict the noise parameters $\rightarrow$ the network recovers this 'system drift'

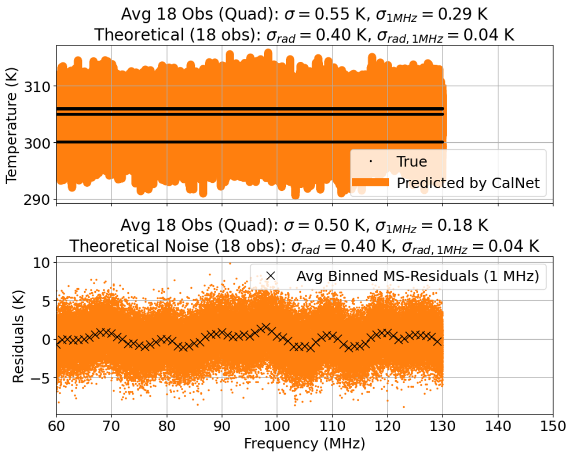
Single night, many observations



Many nights, many observations



Calibration Performance



- Significant improvement when combining many nights
- Calibration down to 0.18K
- Getting close to theoretical noise

Read the paper...

arXiv > astro-ph > arXiv:2504.16791

Search...

Help | Adv

Astrophysics > Instrumentation and Methods for Astrophysics

[Submitted on 23 Apr 2025]

Radiometer Calibration using Machine Learning

S. A. K. Leeney, H. T. J. Bevins, E. de Lera Acedo, W. J. Handley, C. Kirkham, R. S. Patel, J. Zhu, D. Molnar, J. Cumner, D. Anstey, K. Artuc, G. Bernardi, M. Bucher, S. Carey, J. Cavillot, R. Chiello, W. Croukamp, D. I. L. de Villiers, J. A. Ely, A. Fialkov, T. Gessey-Jones, G. Kulkarni, A. Magro, P. D. Meerburg, S. Mittal, J. H. N. Pattison, S. Pegwal, C. M. Pieterse, J. R. Pritchard, E. Puchwein, N. Razavi-Ghods, I. L. V. Roque, A. Saxena, K. H. Scheutwinkel, P. Scott, E. Shen, P. H. Sims, M. Spinelli

Radiometers are crucial instruments in radio astronomy, forming the primary component of nearly all radio telescopes. They measure the intensity of electromagnetic radiation, converting this radiation into electrical signals. A radiometer's primary components are an antenna and a Low Noise Amplifier (LNA), which is the core of the ``receiver'' chain. Instrumental effects introduced by the receiver are typically corrected or removed during calibration. However, impedance mismatches between the antenna and receiver can introduce unwanted signal reflections and distortions. Traditional calibration methods, such as Dicke switching, alternate the receiver input between the antenna and a well-characterised reference source to mitigate errors by comparison. Recent advances in Machine Learning (ML) offer promising alternatives. Neural networks, which are trained using known signal sources, provide a powerful means to model and calibrate complex systems where traditional analytical approaches struggle. These methods are especially relevant for detecting the faint sky-averaged 21-cm signal from atomic hydrogen at high redshifts. This is one of the main challenges in observational Cosmology today. Here, for the first time, we introduce and test a machine learning-based calibration framework capable of achieving the precision required for radiometric experiments aiming to detect the 21-cm line.

Comments: Under peer review for publication in Nature Scientific Reports as part of the Radio Astronomy collection

Subjects: **Instrumentation and Methods for Astrophysics (astro-ph.IM)**; Cosmology and Nongalactic Astrophysics (astro-ph.CO); Artificial Intelligence (cs.AI)

Cite as: arXiv:2504.16791 **[astro-ph.IM]**

(or arXiv:2504.16791v1 **[astro-ph.IM]** for this version)

https://doi.org/10.48550/arXiv.2504.16791

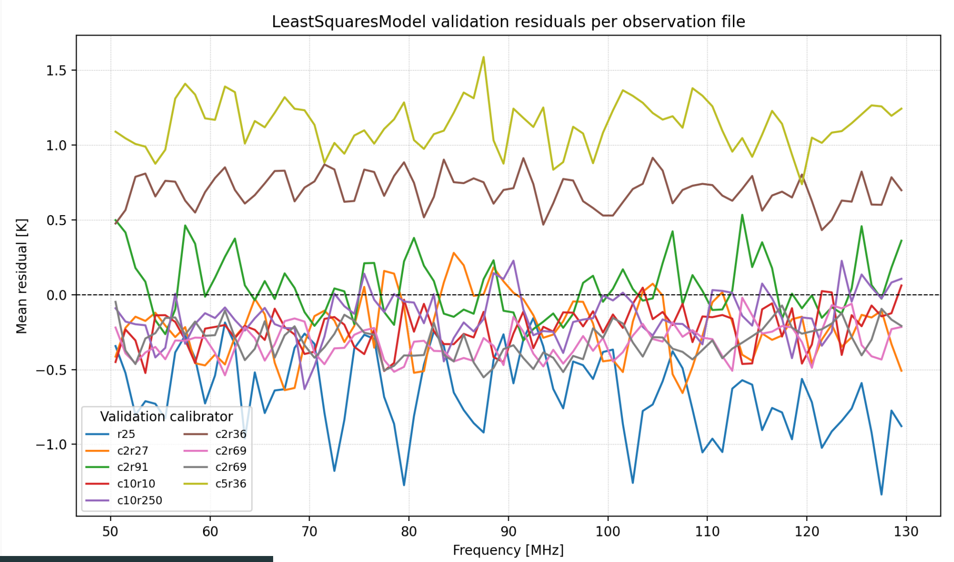
Submission history

From: Samuel Alan Kossoff Leeney [view email]

[v1] Wed, 23 Apr 2025 15:10:25 UTC (25,900 KB)

Neural corrections

Neural LSQ residual diagnostics



Neural pre-corrected LSQ pipeline

1. Neural network predicts additive corrections: $\Delta T = \mathrm{NN}(f, \Gamma)$.
2. Apply the correction to the measured spectrum: $T_{\text{corrected}} = T_{\text{measured}} + \Delta T$.
3. Run a least-squares solve inside the loss: $\theta = \arg\min_\theta \|X\theta - T_{\text{corrected}}\|^2$.
4. Predict the calibrated temperature: $T_{\text{pred}} = X\theta$.
5. Define the loss: $\mathcal{L} = \|T_{\text{pred}} - T_{\text{corrected}}\|^2$.

The LSQ solver is executed *inside* the training loop, so JAX backpropagates through the closed-form solve.

Neural pre-corrected LSQ mathematics

The LSQ design matrix for frequency f and calibrator c is

$$X_c = \begin{bmatrix} x_u & x_c & x_s & x_{NS} & x_L \end{bmatrix},$$

where the noise-wave formulation provides X from the calibrator and receiver S_{11}.

For each frequency we solve

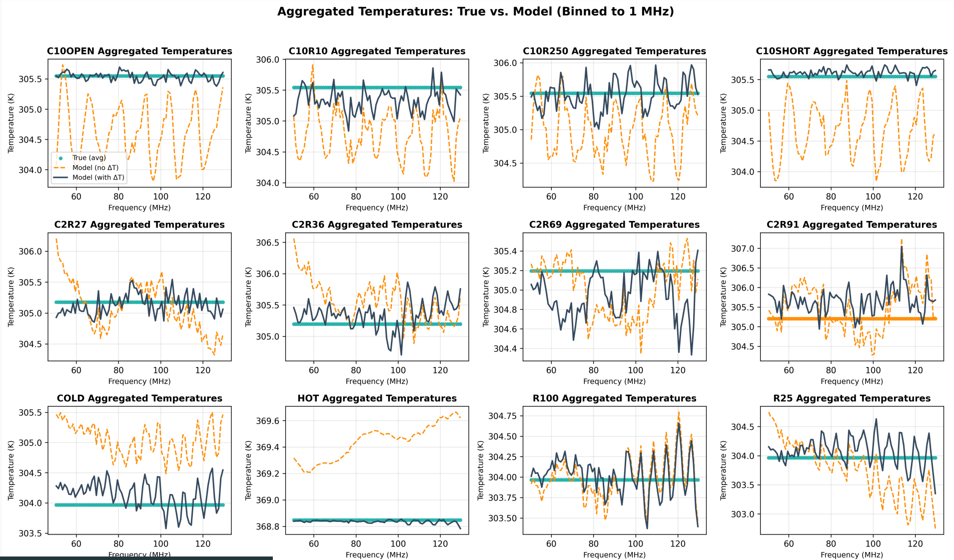
$$\theta(f) = \arg\min_{\theta} \left\| X(f)\theta - \left(T_{\text{measured}} + \mathrm{NN}(f, \Gamma; w) \right) \right\|^2,$$

and the training loss becomes

$$\mathcal{L}(w) = \sum_{f,c} \left[T_{\text{measured}}(c) - \left(X_c(f) \cdot \theta(f) - \mathrm{NN}(f, \Gamma_c; w) \right) \right]^2.$$

Optimising w therefore yields the neural correction weights that minimise the residual between corrected measurements and LSQ predictions.

Aggregated calibrator temperatures



Neural anomaly detection

Bayesian anomaly detection problem

- Each data point carries an anomaly mask $\varepsilon_i \in \{0, 1\}$ with Bernoulli prior $P(\varepsilon_i) = p_i^{\varepsilon_i}(1 - p_i)^{1-\varepsilon_i}$.
- Likelihood blends the normal model $\mathcal{L}_i(\theta)$ and a uniform anomaly floor $\frac{1}{\Delta}$:

$$P(\mathcal{D}, \vec{\varepsilon}|\theta) = \prod_{i=1}^{N} [\mathcal{L}_i(\theta)(1 - p_i)]^{1-\varepsilon_i} \left[\frac{p_i}{\Delta}\right]^{\varepsilon_i}.$$

- Dominant mask approximation yields the rule $\log \mathcal{L}_i + \log \Delta \lessgtr \mathrm{logit}(p_i)$ for deciding whether a point is anomalous.

Hard vs. soft mixture models

Hard mixture:

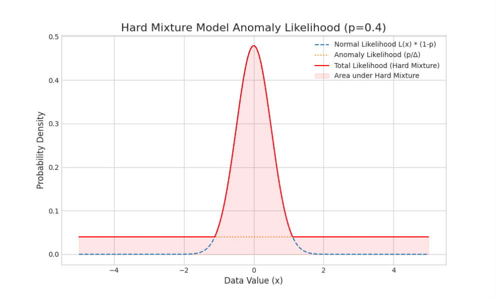
$$\log P(\mathcal{D}|\theta) = \sum_i \max\left\{\log \mathcal{L}_i + \log(1 - p_i), \log p_i - \log \Delta\right\}$$

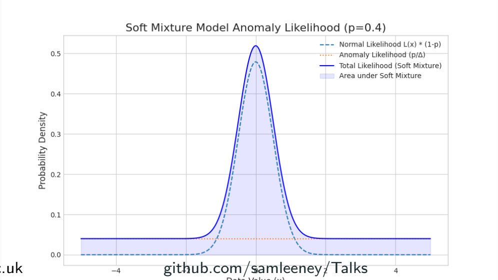
- Fast but clips between hypotheses.
- Sensitive when p_i varies with environment.

Soft mixture:

$$\log P(\mathcal{D}|\theta) = \sum_i \log\left[(1 - p_i)\mathcal{L}_i + \frac{p_i}{\Delta}\right]$$

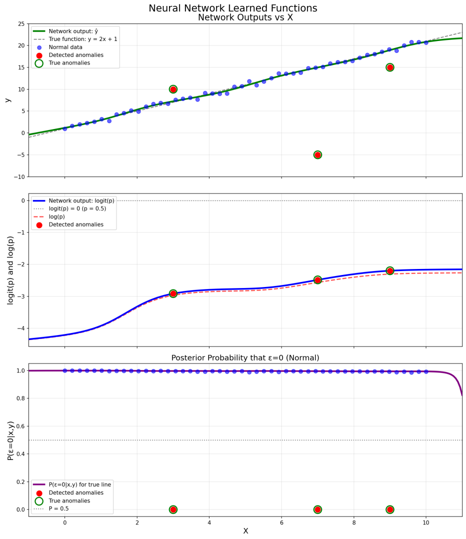
- Smoothly marginalises corruption states.
- Posterior anomaly probability drops below 0.5 when $(1 - p_i)\mathcal{L}_i < \frac{p_i}{\Delta}$.



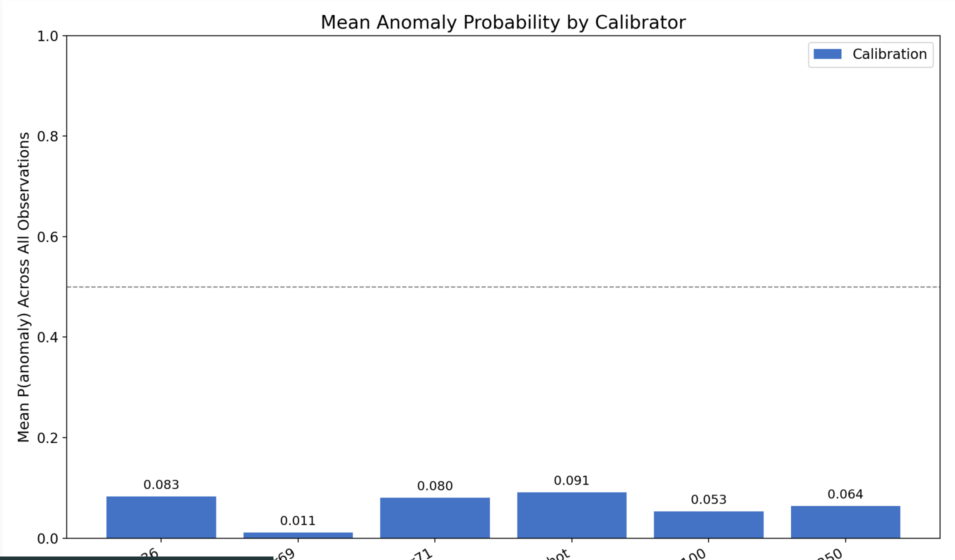


Toy anomaly detection results

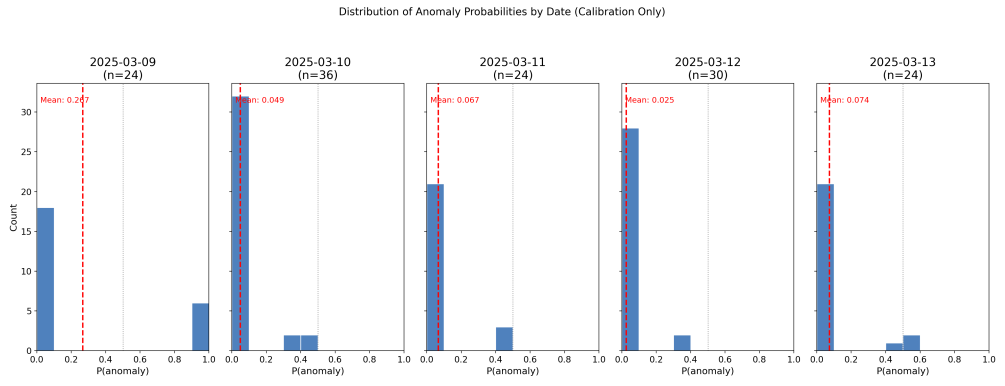
- Synthetic linear regression with Gaussian noise plus injected RFI-like anomalies.
- Network jointly fits $\hat{y}$ and spatial logit(p), achieving perfect 3/3 anomaly recall with zero false positives.
- Posterior $P(\varepsilon = 0|x, y)$ cleanly separates corrupted spectra before calibration.



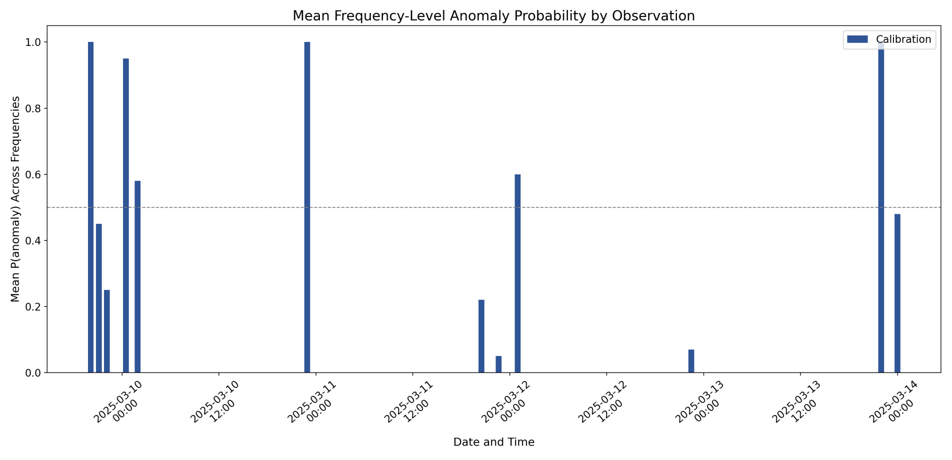
Calibrator-level anomaly statistics



Daily anomaly probability distributions

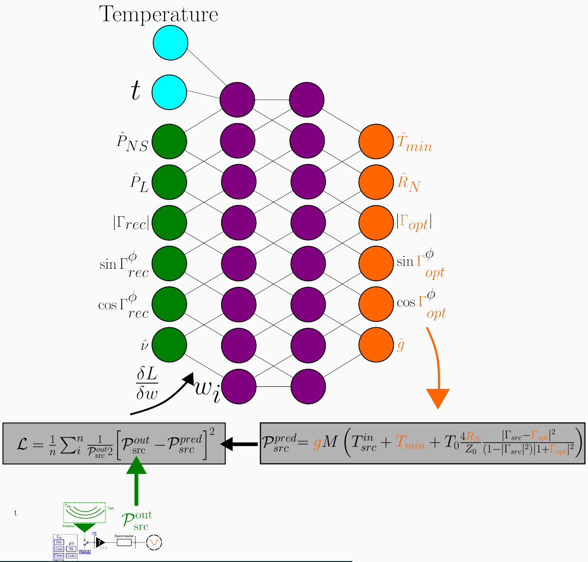


Observation-level anomaly tracking



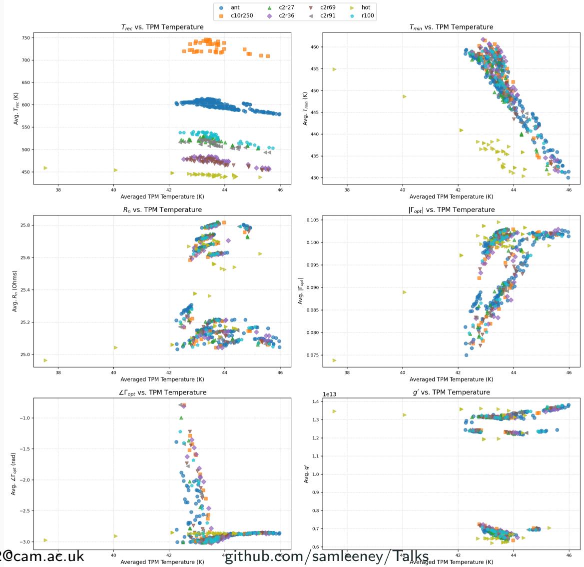
Modeling system drift from other complex system effects

Modeling environmental features



We can train on features that cannot be modeled analytically

Fitting environmental features against fitted noise parameters



Correlate noise parameters with environmental data

sakl2@cam.ac.uk github.com/samleeney/Talks

End to end simulations

End to end simulation

We have shown we can calibrate an internal source, we now test the method on as part of the broader system (simulated).

1) Simulate full receiver chain

↓

2) Simulate realistic antenna with 21cm signal

↓

3) Fit noise parameters to simulated receiver chain

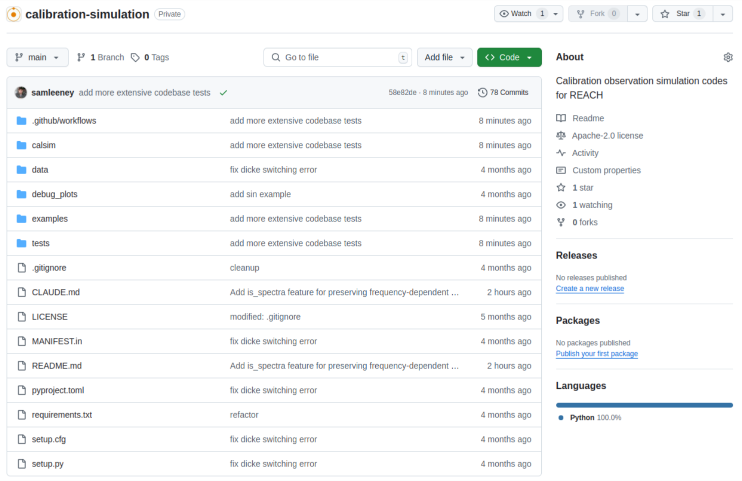
↓

4) Calibrate 'simulated' antenna

↓

5) Fit 21cm signal using REACH pipeline

Simulation pipeline for radiometer calibration



Predicted vs True Antenna Temperature

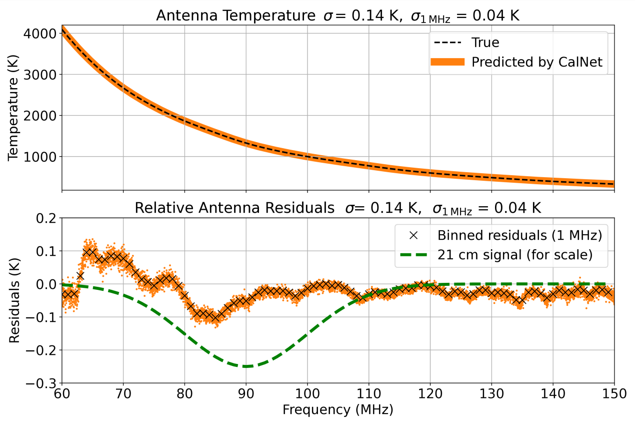


Figure 5: The top panel shows the predicted antenna temperature in orange, with the true temperature overlaid in black dashes

Inferred 21cm Signal

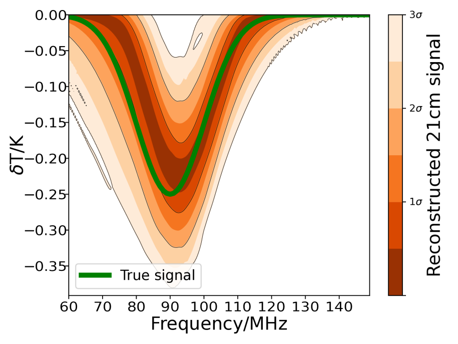


Figure 6: Left: Reconstructed 21cm signal with posterior (orange) and true signal (green)

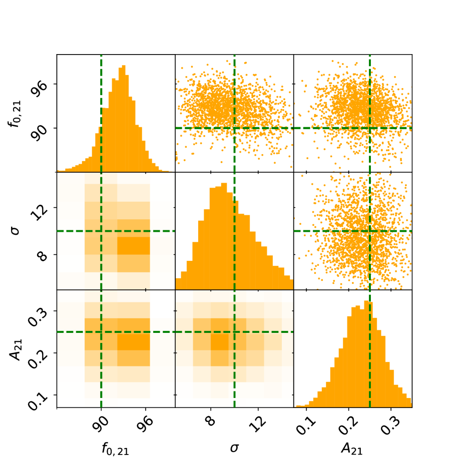


Figure 7: Right: Posterior plots of recovered 21cm signal parameters

Thank you!

Appendix

Calibration Equation

Typically, substitute in the noise wave parameter equation here (gains cancel)

$$T^*_{\mathrm{cal}} = T_{\mathrm{NS}} \frac{P_{\mathrm{cal}} - P_L}{P_{\mathrm{NS}} - P_L} + T_L \tag{4}$$

Make some matching assumptions and re arrange:

$$\begin{aligned} T_{\mathrm{s}} = {} & T_{\mathrm{NS}} \left(\frac{P_{\mathrm{s}} - P_{\mathrm{L}}}{P_{\mathrm{NS}} - P_{\mathrm{L}}} \right) \frac{|1 - \Gamma_{\mathrm{s}}\Gamma_{\mathrm{rec}}|^2}{1 - |\Gamma_{\mathrm{s}}|^2} + T_{\mathrm{L}} \frac{|1 - \Gamma_{\mathrm{s}}\Gamma_{\mathrm{rec}}|^2}{1 - |\Gamma_{\mathrm{s}}|^2} - T_{\mathrm{unc}} \frac{|\Gamma_{\mathrm{s}}|^2}{1 - |\Gamma_{\mathrm{s}}|^2} + \\ & - T_{\cos} \frac{\Re\left(\frac{\Gamma_{\mathrm{s}}}{1 - \Gamma_{\mathrm{s}}\Gamma_{\mathrm{rec}}}\right) |1 - \Gamma_{\mathrm{s}}\Gamma_{\mathrm{rec}}|^2}{(1 - |\Gamma_{\mathrm{s}}|^2)\sqrt{1 - |\Gamma_{\mathrm{rec}}|^2}} - T_{\sin} \frac{\Im\left(\frac{\Gamma_{\mathrm{s}}}{1 - \Gamma_{\mathrm{s}}\Gamma_{\mathrm{rec}}}\right) |1 - \Gamma_{\mathrm{s}}\Gamma_{\mathrm{rec}}|^2}{(1 - |\Gamma_{\mathrm{s}}|^2)\sqrt{1 - |\Gamma_{\mathrm{rec}}|^2}} \end{aligned} \tag{5}$$

Note: We end up with 5 parameters that need to be estimated to calibrate the system.

Calculating the error

By partial derivatives To find the error in T_s, we propagate the errors in Γ_s, Γ_{rec}, P_L, P_{NS}, and P_s:

$$(\Delta T_s)^2 = \left(\frac{\partial T_s}{\partial \Gamma_s}\Delta\Gamma_s\right)^2 + \left(\frac{\partial T_s}{\partial \Gamma_{rec}}\Delta\Gamma_{rec}\right)^2 + \left(\frac{\partial T_s}{\partial P_L}\Delta P_L\right)^2 + \left(\frac{\partial T_s}{\partial P_{NS}}\Delta P_{NS}\right)^2 + \left(\frac{\partial T_s}{\partial P_s}\Delta P_s\right)^2. \quad (10)$$

Calculating the error

$$(\Delta T_s)^2 = \left(\frac{\partial T_s}{\partial \Gamma_s}\Delta\Gamma_s\right)^2 + \left(\frac{\partial T_s}{\partial \Gamma_{rec}}\Delta\Gamma_{rec}\right)^2 + \left(\frac{\partial T_s}{\partial P_L}\Delta P_L\right)^2 + \left(\frac{\partial T_s}{\partial P_{NS}}\Delta P_{NS}\right)^2 + \left(\frac{\partial T_s}{\partial P_s}\Delta P_s\right)^2. \tag{11}$$

$$\frac{\partial T_s}{\partial P_L} = T_{\mathrm{NS}}\frac{|1-\Gamma_{\mathrm{s}}\Gamma_{\mathrm{rec}}|^2}{1-|\Gamma_{\mathrm{s}}|^2}\cdot\frac{P_{\mathrm{s}}-P_{\mathrm{NS}}}{(P_{\mathrm{NS}}-P_{\mathrm{L}})^2}, \tag{12}$$

$$\frac{\partial T_s}{\partial P_{NS}} = -T_{\mathrm{NS}}\frac{|1-\Gamma_{\mathrm{s}}\Gamma_{\mathrm{rec}}|^2}{1-|\Gamma_{\mathrm{s}}|^2}\cdot\frac{P_{\mathrm{s}}-P_{\mathrm{L}}}{(P_{\mathrm{NS}}-P_{\mathrm{L}})^2}, \tag{13}$$

$$\frac{\partial T_s}{\partial P_s} = T_{\mathrm{NS}}\left(\frac{1}{P_{\mathrm{NS}}-P_{\mathrm{L}}}\right)\frac{|1-\Gamma_{\mathrm{s}}\Gamma_{\mathrm{rec}}|^2}{1-|\Gamma_{\mathrm{s}}|^2}. \tag{14}$$

$$\frac{\partial T_s}{\partial \Gamma_{rec}} = \frac{\partial A}{\partial \Gamma_{rec}} + \frac{\partial B}{\partial \Gamma_{rec}} + \frac{\partial D}{\partial \Gamma_{rec}} + \frac{\partial E}{\partial \Gamma_{rec}}. \tag{15}$$

$$\frac{\partial T_s}{\partial \Gamma_s} = \frac{\partial A}{\partial \Gamma_s} + \frac{\partial B}{\partial \Gamma_s} + \frac{\partial C}{\partial \Gamma_s} + \frac{\partial D}{\partial \Gamma_s} + \frac{\partial E}{\partial \Gamma_s}. \tag{16}$$

$$A = T_{\mathrm{NS}}\left(\frac{P_{\mathrm{s}}-P_{\mathrm{L}}}{P_{\mathrm{NS}}-P_{\mathrm{L}}}\right)\frac{|1-\Gamma_{\mathrm{s}}\Gamma_{\mathrm{rec}}|^2}{1-|\Gamma_{\mathrm{s}}|^2}, \tag{17}$$

$$B = T_{\mathrm{L}}\frac{|1-\Gamma_{\mathrm{s}}\Gamma_{\mathrm{rec}}|^2}{1-|\Gamma_{\mathrm{s}}|^2}, \tag{18}$$

$$C = -T_{\mathrm{unc}}\frac{|\Gamma_{\mathrm{s}}|^2}{1-|\Gamma_{\mathrm{s}}|^2}, \tag{19}$$

$$D = -T_{\cos}\frac{\Re\left(\frac{\Gamma_{\mathrm{s}}}{1-\Gamma_{\mathrm{s}}\Gamma_{\mathrm{rec}}}\right)|1-\Gamma_{\mathrm{s}}\Gamma_{\mathrm{rec}}|^2}{(1-|\Gamma_{\mathrm{s}}|^2)\sqrt{1-|\Gamma_{\mathrm{rec}}|^2}}, \tag{20}$$

$$E = -T_{\sin}\frac{\Im\left(\frac{\Gamma_{\mathrm{s}}}{1-\Gamma_{\mathrm{s}}\Gamma_{\mathrm{rec}}}\right)|1-\Gamma_{\mathrm{s}}\Gamma_{\mathrm{rec}}|^2}{(1-|\Gamma_{\mathrm{s}}|^2)\sqrt{1-|\Gamma_{\mathrm{rec}}|^2}}. \tag{21}$$

Calculating the error

Using $T_{\mathbf{NS}} \frac{P_{\mathbf{cal}} - P_L}{P_{\mathbf{NS}} - P_L} + T_L$

$$(\Delta T_s)^2 = \left(\frac{\partial T_s}{\partial P_s}\Delta P_s\right)^2 + \left(\frac{\partial T_s}{\partial P_L}\Delta P_L\right)^2 + \left(\frac{\partial T_s}{\partial P_{NS}}\Delta P_{NS}\right)^2 \tag{22}$$

$$+ \left(\frac{\partial T_s}{\partial \Gamma_s}\Delta \Gamma_s\right)^2 + \left(\frac{\partial T_s}{\partial \Gamma_{rec}}\Delta \Gamma_{rec}\right)^2 . \tag{23}$$

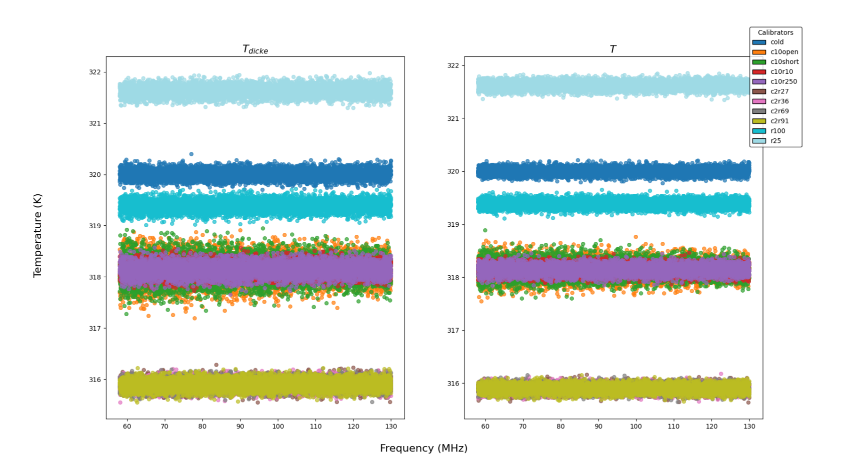
Using noise wave parameters only

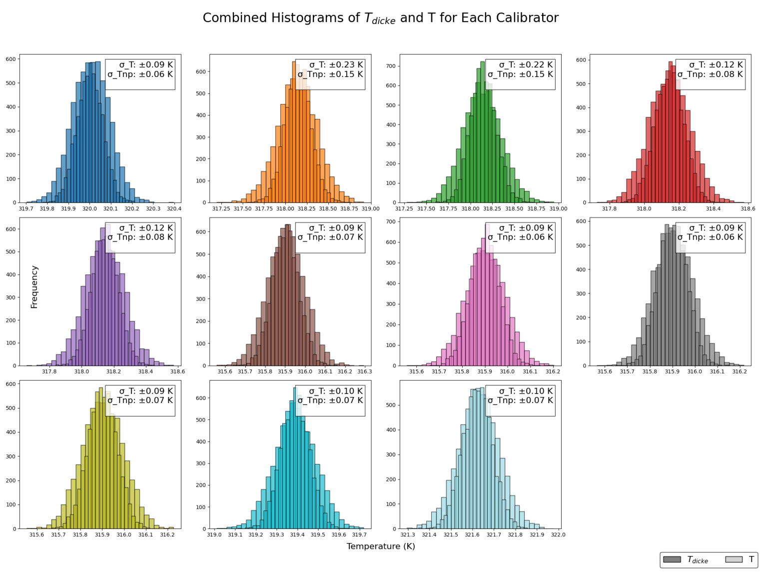
$$(\Delta T_{\mathrm{s}})^2 = \left(\frac{\partial T_{\mathrm{s}}}{\partial P_{\mathrm{s}}}\Delta P_{\mathrm{s}}\right)^2 \tag{24}$$

$$+ \left(\frac{\partial T_{\mathrm{s}}}{\partial \Gamma_{\mathrm{s}}}\Delta \Gamma_{\mathrm{s}}\right)^2 + \left(\frac{\partial T_{\mathrm{s}}}{\partial \Gamma_{\mathrm{rec}}}\Delta \Gamma_{\mathrm{rec}}\right)^2 . \tag{25}$$

Note that this argument applies for both noise parameters and noise wave parameters.

There is more noise when using $T_{\mathrm{NS}}\frac{P_{\mathrm{cal}}-P_L}{P_{\mathrm{NS}}-P_L}+T_L$





Noise amplified by **30%** when using

$$T_{\mathrm{NS}} \frac{P_{\mathrm{cal}} - P_L}{P_{\mathrm{NS}} - P_L} + T_L$$

Why not fit noise (wave) parameters directly?

Noise Parameter Equation:

$$P_{\text{out}}^{\text{src}} = gM\left(T_{\text{in}}^{\text{src}} + T_{\text{min}} + T_0 \frac{4R_N}{Z_0} \frac{|\Gamma_{\text{src}} - \Gamma_{\text{opt}}|^2}{(1 - |\Gamma_{\text{src}}|^2)(1 + |\Gamma_{\text{opt}}|^2)}\right) \tag{26}$$

Noise Wave Equation:

$$\begin{aligned} P_{\text{out}}^{\text{src}} = g\Bigg[& T_0 + T_{\text{unc}}|\Gamma_{\text{s}}|^2 \left|\frac{\sqrt{1 - |\Gamma_{\text{rec}}|^2}}{1 - \Gamma_{\text{s}}\Gamma_{\text{rec}}}\right|^2 \\ & + T_{\text{s}}(1 - |\Gamma_{\text{s}}|^2)\left|\frac{\sqrt{1 - |\Gamma_{\text{rec}}|^2}}{1 - \Gamma_{\text{s}}\Gamma_{\text{rec}}}\right|^2 + T_{\cos}\Re\left(\Gamma_{\text{s}}\frac{\sqrt{1 - |\Gamma_{\text{rec}}|^2}}{1 - \Gamma_{\text{s}}\Gamma_{\text{rec}}}\right) \\ & + T_{\sin}\Im\left(\Gamma_{\text{s}}\frac{\sqrt{1 - |\Gamma_{\text{rec}}|^2}}{1 - \Gamma_{\text{s}}\Gamma_{\text{rec}}}\right)\Bigg] \end{aligned} \tag{27}$$

We still end up with 5 unknowns, as before.